Contemporary Mathematics

Threads Through Group Theory

Persi Diaconis

Abstract. This paper records the path of a letter that Marty Isaacs wrote to a stranger. The tools in the letter are used to illustrate a different way of studying random walk on the Heisenberg group. The author also explains how the letter contributed to the development of super-character theory.

1. Introduction

Marty Isaacs believes in answering questions. They can come from students, colleagues, or perfect strangers. As long as they seem serious, he usually gives it a try. This paper records the path of a letter that Marty wrote to a stranger (me). His letter was useful; it is reproduced in the Appendix and used extensively in Section 3. It led to more correspondence and a growing set of extensions.

Here is some background. I am a mathematical statistician who, for some unknown reason, loves finite group theory. I was trying to make my own peace with a corner of p-group theory called *extra-special p-groups*. These are p-groups G with center $Z(G)$ equal to commutator subgroup G' equal to the cyclic group C_p. I found considerable literature about the groups [2, Sect. 23], [22, Chap. III, Sect. 13], [39, Chap. 4, Sect. 4] but no "stories". Where do these groups come from? Who cares about them, and how do they fit into some larger picture?

My usual way to understand a group is to invent a "natural" random walk and study its rate of convergence to the uniform distribution. This often calls for detailed knowledge of the conjugacy classes, characters, and of the geometry of the Cayley graph underlying the walk. On a whim, I wrote to Marty Isaacs. My letter has not survived, but it probably looked like this.

> Dear Professor Isaacs,
>
> I am a mathematical statistician who is studying a probability problem involving random walk on the extra-special p-groups. For this, I need detailed knowledge of the conjugacy classes and characters. I wonder if you can point me towards what is known? Any other information about these groups would be most welcome. Thank you in advance for your trouble. Sincerely,
>
> Persi Diaconis

2000 *Mathematics Subject Classification.* Primary 60J20.
This work was supported in part by NSF grant DMS 0505673.

Marty's wonderful answer is reproduced in the Appendix. It contains all that I needed to try going forward. I hit some snags, and just recently found a reasonably direct way to use the characters and comparison theory to solve the original problem. Since this combination, characters + comparison, is an absolutely basic approach to studying random walk, I have high hopes that this will be broadly useful.

Section 2 contains needed background on random walk on finite groups. Section 3 works things out for the Heisenberg group of 3×3 uni-upper-triangular matrices with entries in C_p. Section 4 shows how the Heisenberg example and some known results about random walk on Abelian groups gives a complete solution for the extra-special groups. The final section outlines an approach to the open problem of extending the analysis to the group of $n \times n$ uni-upper-triangular matrices with coefficients in C_p. This leads to the study of *super-characters*, a subject developed in later work with Marty.

What is the effect of kindness to strangers? This paper follows one of those threads. I think our subject is woven from these. Marty has spun out hundreds of threads which lead to all corners of group theory. We are in his debt.

2. Random walk on finite groups

Introductions to random walk on finite groups appear in [5, Chapt. 3], [6] and [24]. The Fourier analytic approach based on characters and spherical functions is developed in [4]. Comprehensive surveys are in [20] and [33, 34]. All of these contain pointers to a growing literature.

This section sets up the basic problems and notation. It shows how characters can be used to give bounds for random walks generated by conjugacy classes. This is illustrated for two examples (used later) on C_m. Finally, comparison theory is introduced and used in conjunction with character theory to give bounds on rates of convergence for general walks. These techniques are applied in the sections that follow.

2.1. Random walk. Let G be a finite group and $S = S^{-1}$ a symmetric set of generators. To avoid parity problems, suppose $\mathrm{id} \in S$. The set S may be used to run a random walk. Informally, pick $s_1, s_2, s_3, \ldots$ uniformly at random from S (with replacement). The walk starts at id and proceeds as

$$\mathrm{id}, s_1, s_2 s_1, s_3 s_2 s_1, \ldots.$$

More formally, define

$$Q(g) = \begin{cases} 1/|S| & \text{if } g \in S \\ 0 & \text{otherwise.} \end{cases} \tag{2.1}$$

Then $Q * Q(g) = \sum_{h \in G} Q(h) Q(gh^{-1})$, $Q^{*k}(g) = \sum_h Q(h) Q^{*k-1}(gh^{-1})$. Then $Q^{*k}(g)$ is the chance that the walk is at g after k steps. Denote the uniform distribution by $U(g) = 1/|G|$. Under our conditions, $Q^{*k}(g) \to U(g)$ as $k \to \infty$. The same result holds for any probability distribution Q which is not supported on

a coset of a subgroup. Convergence is measured by total variation distance:

$$\begin{aligned}\left\|Q^{*k}-U\right\| &= \max_{A\subseteq G}\left|Q^{*k}(A)-U(A)\right| \\ &= \frac{1}{2}\sum_{g}\left|Q^{*k}(g)-U(g)\right| \\ &= \frac{1}{2}\sup_{\|f\|_\infty\le 1}\left|Q^{*k}(f)-U(f)\right|.\end{aligned} \tag{2.2}$$

The first equality in (2.2) is a definition. The second equality is proved by noting that the maximum occurs at $A=\{g: Q^{*k}(g)>U(g)\}$. Taking f as the indicator function of this A proves the third equality.

With these definitions, we have a well-posed math problem: Given G, S and $\epsilon>0$, how large a k is required for $\|Q^{*k}-U\|<\epsilon$? The references above contain many examples and techniques for studying this problem. The present paper focuses on analytic techniques involving characters and comparison.

2.2. Character theory. Suppose that $Q(g)$ is a class function, $Q(g)=Q(h^{-1}gh)$. Then, for a character χ, the *Fourier transform* is defined by $\hat{Q}(\chi)=\frac{1}{\chi(1)}\sum_g \chi(g)Q(g)$. The basic upper bound lemma [5, p. 24] gives

$$4\left\|Q^{*k}-U\right\|^2 \le \sum_{\chi\neq 1}\chi^2(1)\left|\hat{Q}(\chi)\right|^{2k}. \tag{2.3}$$

The right side is a sum over non-trivial irreducible characters. It can sometimes be usefully approximated provided a detailed knowledge of the dimensions and other character values are available. This entails some analysis as well. The following simple example gives a picture of the work involved. It is a warm-up for the more difficult Example 2.3 which is used in Section 3 and Section 4.

Example 2.1 (Random walk on C_m). For $m\ge 3$, let C_m be the integers (mod m). Take $S=\{0,1,-1\}$. The characters $\chi_j(k)=e^{2\pi ijk/m}$, $0\le j\le m-1$, are one-dimensional. The Fourier transform is $\hat{Q}(\chi_j)=\frac{1}{3}+\frac{2}{3}\cos\left(\frac{2\pi j}{m}\right)$. The bound in (2.3) becomes

$$4\left\|Q^{*k}-U\right\|^2 \le \sum_{j=1}^{m-1}\left[\frac{1}{3}+\frac{2}{3}\cos\left(\frac{2\pi j}{m}\right)\right]^{2k}. \tag{2.4}$$

The task of bounding the rate of convergence now becomes the analytic problem, How large should k be (as a function of m) so that the right-hand side of (2.4) is small? Because $\cos(-x)=\cos(x)$, the terms for $j\ge(m-1)/2$ equal the terms for $j\le(m-1)/2$. For j with $\cos\left(\frac{2\pi j}{m}\right)\le 0$, the term $|\frac{1}{3}+\frac{2}{3}\cos\left(\frac{2\pi j}{m}\right)|^{2k}\le\left(\frac{1}{3}\right)^{2k}$. It follows that we may bound

$$m\left(\frac{1}{3}\right)^{2k}+2\sum_{j=1}^{(m-1)/4}\left[\frac{1}{3}+\frac{2}{3}\cos\left(\frac{2\pi j}{m}\right)\right]^{2k}.$$

To save work, we will use the inequality $\cos(x)\le e^{-x^2/2}$ for $0\le x\le\pi/2$. Further, $e^{-x^2/2}$ is concave on $[-1,1]$, so $\left[\frac{1}{3}+\frac{2}{3}\cos\left(\frac{2\pi j}{m}\right)\right]\le e^{-\frac{1}{2}\left(\frac{2}{3}\frac{2\pi j}{m}\right)^2}$. Taking $k=cm^2$,

we have

$$4\left\|Q^{*k}-U\right\|^2 \le m\left(\frac{1}{3}\right)^{2cm^2} + 2\sum_{j=1}^{\infty} e^{-c\frac{4}{18}\pi^2 j^2}.$$

Clearly the right side tends to zero, exponentially fast, as c tends to infinity. This shows that $k = cm^2$ steps suffice for convergence. To show that this cannot be substantially improved, use (2.2) with $f(j) = e^{2\pi i j/m}$. Then $U(f) = 0$ and, using $\cos x = 1 - \frac{x^2}{2} + O(x^4)$,

$$\begin{aligned}\left\|Q^{*k}-U\right\| &\ge \frac{1}{2}\left[\frac{1}{3}+\frac{2}{3}\cos\left(\frac{2\pi}{m}\right)\right]^k \\ &= \left[1-\frac{8\pi^2}{9m^2}+O\left(\frac{1}{m^4}\right)\right]^{cm^2} \sim e^{-\frac{8}{9}\pi^2 c}.\end{aligned}$$

This shows that no fixed multiple of m^2 will drive the total variation distance to zero.

Summarizing, we have proved the following.

THEOREM 2.2. *For $m \ge 3$ and $c > 0$, the random walk on C_m generated by $\{0, \pm 1\}$ satisfies*

$$L(c) \le \left\|Q^{*k}-U\right\| \le U(c) \qquad \text{for } k = cm^2$$

with explicit continuous positive functions $L(c)$, $U(c)$, with $U(c) \to 0$ as $c \to \infty$.

The program of proving sharp rates of convergence for random walk on finite groups using characters was first carried out for the random transpositions walk on the symmetric group S_n [13]. See further [4, 5, 8, 23, 28, 32, 35]. It has been successful for random walk on Lie-type groups [1, 16, 17, 19, 25, 26, 29]. It has also been very useful for finite affine groups [3, 20, 21].

It has not been so useful for p-groups, partially because the generating conjugacy classes are often huge (so the walk is random in a few steps and there is no asymptotics to do). Another problem is that the knowledge of conjugacy classes and characters is not available. Before turning to this, we treat a second example (used below) which shows that there is still a lot to do for finite Abelian groups.

EXAMPLE 2.3 (A faster walk on C_m). Take $m = n^2$ and $S = \{0, \pm 1, \pm n\}$. The following argument shows that order cm steps are necessary and sufficient for convergence. A similar argument works for C_p with p prime and generators $\{0, \pm 1 \pm \lfloor\sqrt{p}\rfloor\}$.

The quantity to be bounded is

$$\begin{aligned}(2.5)\quad &\sum_{j=1}^{n^2-1}\left[\frac{1}{5}+\frac{2}{5}\cos\left(\frac{2\pi j}{n^2}\right)+\frac{2}{5}\cos\left(\frac{2\pi j}{n}\right)\right]^{2k} \\ &\qquad = \sum_{j_1,j_2}\left\{\frac{1}{5}+\frac{2}{5}\cos\left(\frac{2\pi(j_1+nj_2)}{n^2}\right)+\frac{2}{5}\cos\left(\frac{2\pi(j_1+nj_2)}{n}\right)\right\}^{2k}.\end{aligned}$$

The sum is over $0 \le j_1, j_2 \le n-1$ with $0 < j_1 + nj_2 \le n^2 - 1$. Assume that $k = cn^2$ throughout. Because $\cos(x) = \cos(-x)$, we may bound twice the sum over $0 < j < (n^2-1)/2$. If $\cos\left(\frac{2\pi j}{n^2}\right) \le 0$, the term in the sum is at most $\left(\frac{3}{5}\right)$ and so,

adding the negligible term $n^2\left(\frac{3}{5}\right)^{2cn^2}$, we may assume $0 \le j \le n^2/4$. The last term is $\frac{2}{5}\cos\left(\frac{2\pi j_1}{n}\right)$. If $j_1 \ge n^{1/4}$, the term in curly brackets can be bounded by

$$\frac{3}{5}+\frac{2}{5}\cos\left(\frac{2\pi j_1}{n}\right) \le e^{-\frac{1}{2}\left(\frac{2}{5}\frac{j_1}{n}\right)^2}.$$

Raising this to the power $2k = 2cn^2$ gives a term at most $e^{-\frac{2c}{25}n^{1/4}}$. Multiplying this by the number of terms (n^2-1) gives something negligible. Hence it may be assumed that $j_1 \le n^{1/4}$ from now on. Similarly, if $j_2 \ge n^{1/4}$, write

$$\cos\left(\frac{2\pi j_1 + nj_2}{n^2}\right) \;=\; \cos\left(\frac{2\pi j_1}{n^2}\right)\cos\left(\frac{2\pi j_2}{n}\right) \;-\; \sin\left(\frac{2\pi j_1}{n^2}\right)\sin\left(\frac{2\pi j_2}{n}\right).$$

Because $j_1 \le n^2/4$, $\sin\left(\frac{2\pi j_1}{n^2}\right)\sin\left(\frac{2\pi j_2}{n}\right) > 0$ and omitting this term increases the summand. Further, $j_1 \le n^{1/4}$ gives $\cos\left(\frac{2\pi j_1}{n^2}\right) > 0$. If $j_2 > n/4$, $\cos\left(\frac{2\pi j_1}{n^2}\right)\cos\left(\frac{2\pi j_2}{n}\right) < 0$ and we may bound all these terms by $n^2\left(\frac{3}{5}\right)^{2cn^2}$ as before. It thus follows that we may assume $n^{1/4} \le j_2 < n/4$. These terms are bounded above by $\left[\frac{3}{5}+\frac{2}{5}\cos\left(\frac{2\pi j_2}{n}\right)\right]^{2cn^2} \le e^{-cn^2\left(\frac{4}{5}\frac{\pi j_2}{n}\right)^2} = e^{-\frac{16\pi^2 c}{25}n^{1/2}}$. The sum of all such terms is again negligible.

Assume finally that $0 \le j_1, j_2 < n^{1/4}$. The corresponding terms are bounded by

$$\begin{aligned}\left[\frac{3}{5}+\frac{2}{5}\cos\left(\frac{2\pi j_1}{n}\right)+\frac{2}{5}\cos\left(\frac{2\pi j_2}{n}\right)\right]^{2cn^2} &\le e^{-cn^2\left(\frac{2}{5}\frac{2\pi j_1}{n}+\frac{2}{5}\frac{2\pi j_2}{n}\right)^2}\\ &\le e^{-\frac{c16\pi^2}{25}\left(j_1^2+j_2^2\right)}.\end{aligned}$$

Since

$$\sum_{\substack{1\le j_1<\infty,\\ 0\le j_2<\infty}} e^{-\frac{c16\pi^2}{25}\left(j_1^2+j_2^2\right)}$$

converges, and tends to zero as $c \to \infty$, we have proved the upper bound in the following theorem. The lower bound is proved by using the test function $e^{2\pi i j/n^2}$ as in Example 2.1.

THEOREM 2.4. *For $n \ge 5$ and $c > 0$, the random walk on C_{n^2} generated by $\{0, \pm 1, \pm n\}$ after $k = cn^2$ steps satisfies*

$$L(c) \le \left\|Q^{*k} - U\right\| \le U(c)$$

with positive continuous functions $L(c), U(c)$ such that $U(c) \to 0$ as $c \to \infty$.

Theorems 2.2 and 2.4 may also be proved from results in [10, 11, 12] since C_m is a nilpotent group of class 2 when the generating set is of bounded size. These papers show that for such walks, order $(\text{diam})^2$ steps are necessary and sufficient for convergence. The diameter of C_m in generators ± 1 is at most $m/2$. The diameter of C_m in generators $\pm 1, \pm\lfloor\sqrt{m}\rfloor$ is of order $\sqrt{m}$.

2.3. Comparison theory. To study more general walks, comparison techniques are useful. These are introduced in [8] for random walk on groups, which is a good source for present purposes. Extensions to more general Markov chains and many examples are in [9, 15]. Suppose Q and $\tilde{Q}$ are symmetric probability measures on G with $\tilde{Q}$ a class function. Think of Q as the measure of interest and

$\tilde{Q}$ as a nice measure about which we know *everything*. The object is to study convolution powers of Q using our knowledge of $\tilde{Q}$. To this end, let E be a symmetric generating set contained in the support of Q. For each g in the support of $\tilde{Q}$, choose and fix a respresentation $g = e_1 e_2 \dots e_l$ with $e_i \in E$. This need not be minimal. Denote $|g| = l$. Let $N(e,g)$ be the number of times $e \in E$ appears in the representation for g. Note that $N(e,g) \le |g|$. Finally, let

$$A = \max_{e \in E} \frac{1}{Q(e)} \sum_g |g| N(e,g) \tilde{Q}(g). \tag{2.6}$$

This measures the average difficulty of expressing the steps of $\tilde{Q}$ by steps of Q.

In [8], a variety of bounds appear relating convergence of Q and $\tilde{Q}$. The following is simple to use in present circumstances.

Proposition 2.5 ([8]). *Let Q and $\tilde{Q}$ be symmetric probabilities on a finite group G with $\tilde{Q}$ constant on conjugacy classes. Let E be a symmetric generating set contained in the support of Q. Then, with A from* (2.3),

$$4\left\|Q^{*k} - U\right\|^2 \le |G|\left(1 - 2Q(\mathrm{id})\right)^{2k} + |G| e^{-k/A} + \sum_{\chi \ne 1} |\chi(1)|^2 \hat{\tilde{Q}}(\chi)^{\lfloor k/A \rfloor}. \tag{2.7}$$

The first two terms on the right side come from parity and negative eigenvalues. They are usually trivial to deal with. The sum on the right is a bound on the convergence of $\tilde{Q}$. Roughly, the proposition shows that if the $\tilde{Q}$ walk is close to random after order l steps, the Q walk is close to random after order Al steps.

A useful corollary comes from taking $\tilde{Q} = U$, the always-available uniform distribution. Then $\hat{\tilde{Q}}(\chi) = 0$ for χ non-trivial irreducible. Bounding A in (2.3) by $(\mathrm{diam})^2 \cdot \max_{g \in \sup Q} 1/Q(g)$ gives the following.

Corollary 2.6. *Let G be a finite group, Q a symmetric probability on G. Let $A^* = \mathrm{diam}^2 \cdot \max_{Q(g)>0} n/Q(g)$ with* diam *the diameter of G using generating set the support of Q. Then, for k with $k/A \ge 1$,*

$$4\left\|Q^{*k} - U\right\|^2 \le |G|\left(1 - 2Q(\mathrm{id})\right)^{2k} + |G| e^{-k/A}.$$

This is a very general bound which is usually useful but not perfect. For Example 2.1 above, with $G = C_m$, Q uniform on $0, \pm 1$, diam $\le m/2$, the bound gives

$$4\left\|Q^{*k} - U\right\|^2 \le m\left(\frac{1}{3}\right)^{2k} + m e^{-4k/m^2}.$$

This shows that k of order $m^2 \log m$ steps suffice for randomness. The character theory estimates reduce this to a k of order m^2 which is best possible. Similar results hold for Example 2.3 and the Heisenberg group of the following section: there is an extra $\log |G|$ factor, and some extra work is required to get rid of it.

Consider the symmetric group S_n with generating set $\{\mathrm{id}, (1,2), C, C^{-1}\}$ with C the n-cycle $(1, 2, \dots, n)$. Then diameter is of order n^2 and the bound shows order $n^5 \log n$ steps suffice. In [8], comparison with the random transpositions walk is used to show that order $n^3 \log n$ steps suffice. Arguments of Wilson [41] show this is the right answer.

Here is one final example showing how Example 2.1 follows from the work done for Example 2.3. Working in C_{n^2}, here $Q(s) = \frac{1}{3}$ for $s \in \{0, \pm 1\}$, $\tilde{Q}(s) = \frac{1}{5}$ for $s \in \{0, \pm 1, \pm n\}$. Since the elements in $\sup(\tilde{Q})$ can be represented using elements in $\sup(Q)$ with length at most n, the comparison constant A is at most $9n^2$. Using (2.7) and the calculations done in Example 2.3, we have, for $k = cn^4$,

$$4 \left\| Q^{*k} - U \right\|^2 \leq n^2 \left(\frac{1}{3} \right)^{2k} + n^2 e^{-k/A} + U(c).$$

This shows that cn^4 steps suffice for uniformity for the walk Q on C_{n^2}.

There are now many examples of comparison arguments in the literature. One that I particularly like uses random walk on the hyperoctrahedral group to analyze a natural problem about mutations of DNA [36].

3. The Heisenberg group

Let $H(m^3)$ be the group of 3×3 uni-upper-triangular matrices with entries taken mod (m). These are denoted

$$\begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \quad \text{or} \quad [x, y; z], \qquad x, y, z \in C_m.$$

A natural walk on $H(m^3)$ is generated by

$$S = \{[0, 0; 0], [\pm 1, 0; 0], 0, \pm 1; 0]\}. \tag{3.1}$$

The associated walk amounts to adding or subtracting a randomly chosen row to the row above, or doing nothing. This walk was introduced by Zack [43] who gives a connection to computer generation of random numbers. It has been solved by using the *geometric theory of Markov chains* in joint work with Laurent Saloff-Coste [10, 11, 12]. These papers use three different approaches; the first uses polynomial growth of the generating set in (3.1) (technically, a condition called *moderate growth*). The second paper uses Nash inequalities and eigenvalue estimates. The third lifts the walk to the free nilpotent group of class 2 on two generators, uses a central limit theorem of Hebisch–Saloff-Coste for the lifted walk, and then a Harnack inequality to transfer back to the finite group.

I have been frustrated that I could not get the right convergence rate using character theory and comparison. The main new result in this paper is a method for doing this. Here is that result.

THEOREM 3.1. *For $m \geq 3$, $c > 0$, $k = cm^2$, the random walk on the Heisenberg group $H(m^3)$ with generating set* (3.1) *satisfies*

$$L(c) \leq \left\| Q^{*k} - U \right\| \leq U(c)$$

for explicit positive continuous functions $L(c)$, $U(c)$, with $U(c) \to \infty$ as $c \to \infty$.

The proof follows from knowledge of the conjugacy, characters, and a diameter bound for an enlarged generating set. These will be developed first. To begin, it is easy to check that the center $Z(H(m^3))$ and commutator $H'(m^3)$ are isomorphic to $C_m = \{[0, 0; z],\, z \in C_m\}$. Marty's letter includes a self-contained proof of the following classical facts.

PROPOSITION 3.2. *In $H(m^3)$*

(1) *The conjugacy classes are*
- *the m-elements of the center* $[0,0;z]$
- *the* m^2-1 *classes* $\{[x,y;*],\ *\in C_m\}$

(2) *The characters are*
- *the* m^2 *linear characters* $\chi_{ab}([x,y;z]) = e^{2\pi i(ax+by)/m},\ a,b\in C_m$
- *the* $m-1$ *characters of degree* m, *for all* $0\neq c\in C_m$

$$\chi_c\left([x,y;z]\right) = \begin{cases} me^{2\pi iz/m} & \text{if } x=y=0 \\ 0 & \text{otherwise.}\end{cases}$$

The natural conjugacy-invariant walk results from simply conjugating the generators (3.1). This does not give a good result. The associated walk still takes order m^2 steps to become random. The added factor from comparison gives a bound of order m^3 for the original walk. The new idea in this section is to use an enlarged set of generators. These are described next.

PROPOSITION 3.3. *Let* $\alpha = \lfloor\sqrt{m}\rfloor$. *The diameter of* $H(m^3)$ *in the generating set*

$$[0,\pm1;0],[0,\pm\alpha;0][\pm1,0;0],[\pm\alpha,0;0] \tag{3.2}$$

is at most $4\sqrt{m}$.

PROOF. Using $\pm1,\pm\alpha$, the diameter of C_m is at most $\sqrt{m}$: we may add multiples of α to get within $\sqrt{m}$ of any number and then fill in the remainder with ±1. Now note that $[-1,y,0][1,y,0]=[0,0,-y]$. It follows that any central element can be written as a product of at most $2\sqrt{m}$ generators. This gives a diameter bound of $4\sqrt{m}$ for a general element of $H(m^3)$. □

PROOF OF THEOREM 3.1. For simplicity, take $m=n^2$ and $\alpha=n$ as in Example 2.3. Let $\tilde{Q}$ be the measure putting mass $1/9$ on the identity and measure $1/(9m)$ on each of the conjugates of the generators in (3.2). Using Proposition 3.2, the quantity A of (2.3) is bounded by $80n^2$ ($80=5\cdot16$) and $N(e,y)\le|y|$). The bound from Proposition 2.5 is

$$4\left\|Q^{*k}-U\right\|^2 \le n^6\left(\frac{3}{5}\right)^{2k} + n^6e^{-k/A} + n^2\left(n^2-1\right)\left(\frac{1}{9}\right)^{\lfloor k/A\rfloor}$$
$$+\sum_{a,b\neq0,0}\left\{\frac{1}{9}+\frac{2}{9}\cos\left(\frac{2\pi a}{n^2}\right)+\frac{2}{9}\cos\left(\frac{2\pi a}{n}\right)+\frac{2}{9}\cos\left(\frac{2\pi b}{n^2}\right)+\cos\left(\frac{2\pi b}{n}\right)\right\}^{\lfloor k/A\rfloor}.$$

The first two terms of the bound come from the proposition. The third term comes from the Fourier transform of $\tilde{Q}$ at the n^2-1 of dimension n^2. The sum comes from the Fourier transform at the non-trivial linear characters. Clearly the first three terms are negligible when k is of order n^4.

To bound the sum, note that if any of the four cosine terms is negative, the term inside the curly brackets is bounded in absolute value by $1/9$. All of the terms can be bounded by $m^4\left(\frac{1}{9}\right)^{\lfloor k/A\rfloor}$. Using also $\cos(-x)=\cos(x)$, we must bound

$$\sum_{\substack{0\le a,b\le n^2/4\\ a,b\neq0,0}}\left\{\frac{1}{9}+\frac{2}{9}\cos\left(\frac{2\pi a}{n^2}\right)+\frac{2}{9}\cos\left(\frac{2\pi a}{n}\right)+\frac{2}{9}\cos\left(\frac{2\pi b}{n^2}\right)+\cos\left(\frac{2\pi b}{n}\right)\right\}^{\lfloor k/A\rfloor}.$$

Now, the bounding closely follows Example 2.3. Write $a = a_1 + na_2,\ b = b_1 + nb_2$ for $0 \le a_1, a_2, b_1, b_2 \le n-1$. Use $\cos(x) \le e^{-x^2/2}$ for $0 \le x \le \pi/2$ and properties of cosine to show that terms with any of $a_1, a_2, b_1, b_2 \ge n^{1/4}$ are negligible. Finally, if $a_1, a_2, b_1, b_2 \le n^{1/4}$, the sum is bounded above by

$$D_1 \sum e^{-cD_2\left(d_1^2 + \cdots + d_8^2\right)}$$

for positive constants D_i with the sum over $0 \le d_1, d_2, \ldots, d_8 < \infty$ with $d_1 > 0$. This tends to zero with c tending to infinity. Further details are omitted. □

4. Extra-special p-groups

For p prime, a p-group with center isomorphic to commutator isomorphic to C_p is called *extra-special*. Such a group turns out to have order p^{2n+1} for some $n \ge 1$ and lies in one of two non-isomorphic families.

$$\begin{aligned} H\left(p^{2n+1}\right) = \big\langle x_1, y_1, \ldots, x_n, y_n \big| &[x_i, x_j] = [y_i, y_j] = 1, [x_i, y_j] = 1 \text{ for } i \ne j, \\ &[x_i, y_i] = z \text{ and } x_i^p = y_i^p = z^p = 1 \text{ for } z \in Z\left(H\left(p^{2n+1}\right)\right)\big\rangle \\ M\left(p^{2n+1}\right) = \big\langle x_1, y_1, \ldots, x_n, y_n \big| &[x_i, x_j] = [y_i, y_j] = 1, [x_i, y_j] = 1 \text{ for } i \ne j, \\ &[x_i, y_i] = z \text{ and } y_i^p = z, x_i^p = z^p = 1 \text{ for } z \in Z\left(M\left(p^{2n+1}\right)\right)\big\rangle. \end{aligned}$$

The natural walks choose a generator (or the identity) uniformly. Call the associated measures Q_H and Q_M. Sharp results for these walks were obtained by Richard Stong [37]. Here is one of his results.

THEOREM 4.1 (Stong). *Let Q be the natural random walk on $H(p^{2n+1})$ or $M(p^{2n+1})$.*

- *Suppose $n \to \infty$. Let $k = 2n\log(2n)/(1-\cos(2\pi/p)) + cp^2 n$. Then, there are positive constants C_1, C_2, C_3, C_4 such that*
$$e^{-cC_1} / \left(C_2 + e^{-cC_1}\right) \le \left\|Q^{*k} - U\right\| \le C_3 e^{-cC_4}.$$
- *If n is fixed (or bounded), $p \to \infty$ and $k = cp^2$, there are positive constants $B_i(n)$ such that*
$$e^{-cB_1} / \left(B_2 + e^{-cB_1}\right) \le \left\|Q^{*k} - U\right\| \le B_3 e^{-cB_4}.$$

This gives a different proof of the results in Section 3. For n large, it shows that there is a cut-off in convergence to stationarity. The proof is a completely novel set of techniques involving a decomposition of the transition matrix into blocks which can themselves be interpreted as "twisted" random walks generated by a signed measure. The results do *not* follow from the geometric techniques and we failed in a direct Fourier attack. The following elementary argument works for both series but is presented for the Heisenberg group only.

Let $H(m^{2n+1})$ be the $n+2 \times n+2$ uni-upper-triangular matrices with non-zero entries only in the top row or last column. Let Q be defined by choosing an entry $(1,i), (n+2,i)$ uniformly, $2 \le i \le n-1$, and changing that entry by adding $0, \pm 1$ (chosen uniformly). Despite appearances, this is the random walk above on $H(m^{2n+1})$. The study of the walk generated by Q is carried out by combining the results of Section 3 (for $H(p^3)$) with known results for a random walk on the Abelian group C_m^{2n} generated by choosing a coordinate uniformly and changing it by adding $0, \pm 1$ with probabiltiy $1/3$. In [8, Sect. 6], it is shown that order $m^2 n \log n$ steps

are necessary and sufficient for this walk to be close to random in total variation. The following proposition allows combining results.

PROPOSITION 4.2. *Let μ and ν be probabilities on a finite set $\mathcal{X}$. Let $T : \mathcal{X} \to \mathcal{Y}$ be given and suppose that*

(1) *For some $\epsilon > 0, A \subseteq \mathcal{X}$ and all $t \in \mathcal{Y}$, $|\mu(A|T = t) - \nu(A|T = t)| < \epsilon$*
(2) $\sum_t |\mu(T = t) - \nu(T = t)| < \delta$

Then

$$|\mu(A) - \nu(A)| < \epsilon + \delta.$$

PROOF. Write $\mu^t(A) = \mu(A|T = t)$. Then

$$\begin{aligned}\mu(A) &= \sum_t \mu^t(A) \cdot \mu(t) \\ &= \sum_t \left(\mu^t(A) - \nu^t(A)\right) \mu(t) + \sum_t \nu^t(A) \left(\mu(t) - \nu(t)\right) + \nu(A).\end{aligned}$$

It follows that $|\mu(A) - \nu(A)| \leq \epsilon + \delta$. □

BOUNDING THE WALK ON $H(m^{2n+1})$. Write the steps of the walk as $g_1, g_2, \ldots, g_k$ where g_i are independent and identically distributed from Q on $H(m^{2n+1})$. Divide the steps into two types as they involve the $(1, n+1)$ or $(n+2, 2)$ coordinates (type I) or not (type II). Steps of type I generate a random walk on the subgroup $H(m^3)$ spanned by coordinates $(1, n+1), (1, n+2), (n+2, 2)$. Steps of type II generate a random walk on the subgroup with pattern

$$\begin{pmatrix} 1 & * & \cdots & \cdots & * & 0 & * \\ & 1 & 0 & \vdots & \vdots & \vdots & 0 \\ & & 1 & \vdots & \vdots & \vdots & * \\ & & & \ddots & \vdots & \vdots & \vdots \\ & & & & \ddots & \vdots & * \\ & & & & & & 1 \end{pmatrix}$$

Let B be the event that, with $k = cm^2 n \log n$, the first k steps produce at least $\frac{c}{2}m^2$ steps of type I and $\frac{c}{2}m^2 n \log n$ steps of type II. By elementary large deviations bounds, this is an event of probability $1 - f(m, n, c)$ with $f(m, n, c) \to 0$ as $c \to \infty$. On the event B, the marginal distribution of coordinates in positions $(1, i)$ and $(n+2, i+1)$, $2 \leq i \leq n$, are close to uniformly distributed by [8, Sect. 6]. Further, on B, the distribution of coordinates $(1, n+1), (1, n+2), (n+2, 2)$ are close to uniformly distributed conditional on coordinates $(1, i), (n+2, i+1)$, $2 \leq i \leq n$, uniformly in the conditioning variables. Thus, Proposition 4.2 with $\mu(x) = Q^{*k}(x|B)$ and $\nu(x) = U$ applies. The final result follows from $Q^{*k}(A) = Q^{*k}(A|B)Q^{*k}(B) + Q^{*k}(A|B^C)Q^{*k}(B^C)$. □

5. The full upper-triangular group

A problem extending the study of the groups $H(m^{2n+1})$ is the natural walk on the group $U_n(C_m)$, $n \times n$ uni-upper-triangular matrices with entries mod m. The walk can be described as: Pick row j, $2 \leq j \leq n$, uniformly at random

and add or subtract it to the row above it (or do nothing) with probability 1/3. The comparison approach to this walk, begun in [1] calls for the characters and conjugacy classes of $U_n(C_m)$. This is a well known wild problem and proveably intractable. In a series of papers, Carlos Andre followed by Ning Yan [42] found that by lumping together certain conjugacy classes (into super-classes) and taking sums of certain irreducibles (giving *super-characters*), an elegant theory remained where everything is explicitly computable. It turns out that the conjugacy classes containing the original generators are already super-classes and the cruder theory is all that is needed. This gave useful if not perfect results and improvements have recently emerged. See [1] for details and references to papers by Andre. Recent papers containing improved results are [31, 38]. It seems possible that the new idea of Section 3 can be applied here to improve the comparison bound.

The super-character theory is so elegant it cries out for generalization. I noticed that Marty was working on a class of p-groups called algebra groups, and together we extended things to an elegant super-character theory for these [7]. Nat Thiem and I [14] worked out quite explicit formulae for algebra groups and Thiem–Venkateswaran [40] and Marberg–Thiem [27] have begun to develop restriction and induction formulae for subgroups of $U_n(\mathbb{F}_q)$. The theory strongly reminds me of the combinatorial representation theory of the symmetric group. For S_n, characters are indexed by partitions. For $U_n(\mathbb{F}_q)$, characters are indexed by set partitions. There is very active work by a group of us that holds real promise. It is the subject of an A.I.M. conference set for May, 2010.

Marty has had two Ph.D. students who developed super-character theory in other directions. Anders Hendrickson [18] has begun the classification of Abelian super-character theories. Benjamin Otto has begun a detailed comparison of the super-characters and a collection of fascinating class functions studied by Krylov on $U_n(\mathbb{F}_q)$ [30]. Carlos Andre and others have also made progress.

All of this work can be traced back to Marty's letter. We are still meeting, writing, and following the thread together. I am truly thankful.

Appendix

Letter from Marty Isaacs, January 1994:

> Dear Persi,
>
> The extra-special groups come up all the time in my part of group theory, and so I know them well. I'll tell you what strikes me as the main facts about these groups, and if there is more you would like to know, just ask.
>
> First, an "official" definition. A p-group P is *special* if $P' = Z(P)$ is elementary Abelian. (Note that this implies that the Frattini subgroup is central since $[x^p, y] = [x, y]^p = 1$. It follows that $Z(P) = P'$ is equal to the Frattini subgroup.) The group P is *extra-special* if it is special, and in addition $|Z(P)| = p$.
>
> Now assume P is extra-special and let $Z = Z(P)$. If x is a non-central element of P then its class $\mathrm{cl}(x)$ is contained in Zx since the factor group P/Z is Abelian. Thus $|\mathrm{cl}(x) \leq p$ and it follows that $|\mathrm{cl}(x)| = p$ and $\mathrm{cl}(x) = Zx$. The classes of P are thus exactly the co-sets of Z other than Z itself, together with the p elements of Z as singleton classes.

If we write $|P/Z| = p^m$, we have counted $p^m - 1 + p$ classes. Also, since P/Z is Abelian, there are exactly p^m irreducible characters of degree 1. It follows that there are exactly $p-1$ nonlinear irreducible characters. Since $|P|$ is the sum of the squares of the degrees of all irreducible characters, it follows that the sum of the squares of the $p-1$ nonlinear irreducible characters is $|P| - p^m = p^m(p-1)$. The average of the squares of the degrees of these characters is p^m. If any of these degree squares is above average, it is at least p^{m+1} since it must be a p-power. This is too big, however, since the sum of all $p-1$ degree squares is only $p^m(p-1) < p^{m+1}$. It follows that all of the degree squares are equal to p^m exactly, and in particular, m must be even. We write now $m = 2n$ so $|P| = p^{2n+1}$.

Now let χ be a nonlinear irreducible character of P. Then $\chi(1) = p^n$ and the sum of $|\chi(z)|^2$ for $z \in Z$ is $p^{2n+1} = |P|$. Since the sum of $|\chi(x)|^2$ over the whole group also equals $|P|$, we deduce that χ vanishes on all elements of $P - Z$.

To completely determine χ, we need to evaluate it on elements of Z. Since Z is central, we know that the restriction of χ to Z must be a multiple of a linear character λ of Z. This restriction χ_Z is thus equal to $p^n\lambda$. Moreover, λ can't be trivial because Z is not in the kernel of χ, as P/Z is Abelian. Now different nonlinear irreducible characters of P (recall: there are $p-1$ of these) all vanish off of Z and so they differ on Z. It follows that their restrictions to Z are exactly the characters $p^n\lambda$, as λ runs over the nonprincipal linear characters of Z. We can thus label the nonlinear characters of P as χ_λ, where λ runs over the $p-1$ nontrivial linear characters of Z. A complete description of χ_λ is that its value on x in $P - Z$ is zero and its value on z in Z is $p^n\lambda(z)$.

Of course, the linear characters of P are really just the linear characters of an elementary Abelian p-group of order p^n. I doubt that there is anything I can tell you about those that you don't already know. Observe that the character table of P is completely determined without knowing the isomorphism type.

Let me say a bit about isomorphism type, without being as detailed as in my derivation of the character theory. First, a general fact: if x, y in P don't commute, then $X =< x, y >$ is a little extra-special subgroup of order p^3. (And you know that there are exactly two isomorphism types of extra-special groups of order p^3 for each prime. For $p = 2$ these are D_8 and Q_8 and for $p > 2$, one has exponent p and one has exponent p^2.) Suppose $X < P$ and let Y be the centralizer in P of X. The fact is that $XY = P$ and X intersects Y at Z. Also, Y is an extra-special group. It follows that P is the central product of X and Y. (This is the factor group of the direct product that identifies the centers.)

Repeating this process of splitting off the small group X, we see that P can be written as the central product of n extra-special

groups of order p^3. In particular, if P has exponent P, then all n pieces have exponent p and P is uniquely determined.

Let's investigate the central product of two extra-special groups of order p^3. First assume $p > 2$. One can check that if A and B have exponent p^2, then in their central product AB one can find two noncommuting elements of order p, and these generate (since $p > 2$) an extra-special group X of exponent p. It follows that the central product of two exponent p^2 groups is the same as the central product of one of exponent p and one of exponent p^2. One thus never needs more than one exponent p^2 group to construct P and hence there are two types of P up to isomorphism, one of exponent p and one of exponent p^2.

For $p = 2$, the amazing fact is that the central product of two D_8s is isomorphic to the central product of two Q_8s. It follows that one never needs to use more than one Q_8, and this yields at most two extra-special 2-groups of any given order. It is a fact (but I don't see a quick proof) that there actually are two different groups for any given order.

There is lots more known. There is information about subgroups and about automorphisms, for example, but I'll stop now. If there is anything else you would like to know, please ask.

Marty

References

1. Ery Arias-Castro, Persi Diaconis, and Richard Stanley, *A super-class walk on upper-triangular matrices*, J. Algebra **278** (2004), no. 2, 739–765. MR MR2071663 (2005f:60101)
2. M. Aschbacher, *Finite Group Theory*, second ed., Cambridge Studies in Advanced Mathematics, vol. 10, Cambridge University Press, Cambridge, 2000. MR MR1777008 (2001c:20001)
3. Claudio Asci, *Generating uniform random vectors in $\mathbb{Z}_p^h$*, J. Theoret. Probab. **22** (2009), 791–809.
4. Tullio Ceccherini-Silberstein, Fabio Scarabotti, and Filippo Tolli, *Harmonic Analysis on Finite Groups*, Cambridge Studies in Advanced Mathematics, vol. 108, Cambridge University Press, Cambridge, 2008, Representation theory, Gelfand pairs and Markov chains. MR MR2389056
5. Persi Diaconis, *Group representations in probability and statistics*, Institute of Mathematical Statistics Lecture Notes—Monograph Series, 11, Institute of Mathematical Statistics, Hayward, CA, 1988. MR MR964069 (90a:60001)
6. ———, *Random walks on groups: characters and geometry*, Groups St. Andrews 2001 in Oxford. Vol. I, London Math. Soc. Lecture Note Ser., vol. 304, Cambridge Univ. Press, Cambridge, 2003, pp. 120–142. MR MR2051523 (2005c:20109)
7. Persi Diaconis and I. M. Isaacs, *Supercharacters and superclasses for algebra groups*, Trans. Amer. Math. Soc. **360** (2008), no. 5, 2359–2392. MR MR2373317
8. Persi Diaconis and Laurent Saloff-Coste, *Comparison techniques for random walk on finite groups*, Ann. Probab. **21** (1993), no. 4, 2131–2156. MR MR1245303 (95a:60009)

9. ———, *Comparison theorems for reversible Markov chains*, Ann. Appl. Probab. **3** (1993), no. 3, 696–730. MR MR1233621 (94i:60074)
10. ———, *Moderate growth and random walk on finite groups*, Geom. Funct. Anal. **4** (1994), no. 1, 1–36. MR MR1254308 (95d:60118)
11. ———, *An application of Harnack inequalities to random walk on nilpotent quotients*, Proceedings of the Conference in Honor of Jean-Pierre Kahane (Orsay, 1993), no. Special Issue, 1995, pp. 189–207. MR MR1364885 (97d:60113)
12. ———, *Nash inequalities for finite Markov chains*, J. Theoret. Probab. **9** (1996), no. 2, 459–510. MR MR1385408 (97d:60114)
13. Persi Diaconis and Mehrdad Shahshahani, *Generating a random permutation with random transpositions*, Z. Wahrsch. Verw. Gebiete **57** (1981), no. 2, 159–179. MR MR626813 (82h:60024)
14. Persi Diaconis and Nathaniel Thiem, *Supercharacter formulas for pattern groups*, Trans. Amer. Math. Soc. **361** (2009), no. 7, 3501–3533. MR MR2491890
15. Martin Dyer, Leslie Ann Goldberg, Mark Jerrum, and Russell Martin, *Markov chain comparison*, Probab. Surv. **3** (2006), 89–111 (electronic). MR MR2216963 (2007d:60042)
16. David Gluck, *Characters and random walks on finite classical groups*, Adv. Math. **129** (1997), no. 1, 46–72. MR MR1458412 (98g:20027)
17. ———, *First hitting times for some random walks on finite groups*, J. Theoret. Probab. **12** (1999), no. 3, 739–755. MR MR1702891 (2000i:60007)
18. Anders O. F. Hendrickson, *Supercharacter theories of cyclic p-groups*, Ph.D. thesis, Dept. of Mathematics, University of Wisconsin, 2009.
19. Martin Hildebrand, *Generating random elements in* $\mathrm{SL}_n(\mathbb{F}_q)$ *by random transvections*, J. Algebraic Combin. **1** (1992), no. 2, 133–150. MR MR1226348 (94k:60104)
20. ———, *A survey of results on random random walks on finite groups*, Probab. Surv. **2** (2005), 33–63 (electronic). MR MR2121795 (2006a:60010)
21. ———, *A lower bound for the Chung–Diaconis–Graham random process*, Proc. Amer. Math. Soc. **137** (2009), no. 4, 1479–1487. MR MR2465674
22. B. Huppert, *Endliche Gruppen. I*, Die Grundlehren der Mathematischen Wissenschaften, Band 134, Springer-Verlag, Berlin, 1967. MR MR0224703 (37 #302)
23. Michael Larsen and Aner Shalev, *Characters of symmetric groups: sharp bounds and applications*, Invent. Math. **174** (2008), no. 3, 645–687. MR MR2453603
24. David A. Levin, Yuval Peres, and Elizabeth L. Wilmer, *Markov chains and mixing times*, American Mathematical Society, Providence, RI, 2009, With a chapter by James G. Propp and David B. Wilson. MR MR2466937
25. Martin W. Liebeck and Aner Shalev, *Diameters of finite simple groups: sharp bounds and applications*, Ann. of Math. (2) **154** (2001), no. 2, 383–406. MR MR1865975 (2002m:20029)
26. ———, *Character degrees and random walks in finite groups of Lie type*, Proc. London Math. Soc. (3) **90** (2005), no. 1, 61–86. MR MR2107038 (2006h:20016)
27. Eric Marberg and Nathaniel Thiem, *Superinduction for pattern groups*, J. Algebra **321** (2009), no. 12, 3681–3703. MR MR2517809
28. Thomas W. Müller and Jan-Christoph Schlage-Puchta, *Character theory of symmetric groups, subgroup growth of Fuchsian groups, and random walks*, Adv. Math. **213** (2007), no. 2, 919–982. MR MR2332616 (2008k:20024)

29. R. Nanayakkara, *Diameters and mixing times of classical groups*, Ph.D. thesis, Dept. of Mathematics, Imperial College, London, 2001.
30. Benjamin Otto, *Super-characters for algebra groups*, Ph.D. thesis, Dept. of Mathematics, University of Wisconsin, 2009.
31. Igor Pak, *Two random walks on upper triangular matrices*, J. Theoret. Probab. **13** (2000), no. 4, 1083–1100. MR MR1820503 (2001m:60018)
32. Sandrine Roussel, *Phénomène de cutoff pour certaines marches aléatoires sur le groupe symétrique*, Colloq. Math. **86** (2000), no. 1, 111–135. MR MR1799892 (2001m:60019)
33. Laurent Saloff-Coste, *Probability on groups: random walks and invariant diffusions*, Notices Amer. Math. Soc. **48** (2001), no. 9, 968–977. MR MR1854532 (2003g:60011)
34. ———, *Random walks on finite groups*, Probability on discrete structures, Encyclopaedia Math. Sci., vol. 110, Springer, Berlin, 2004, pp. 263–346. MR MR2023654 (2004k:60133)
35. Laurent Saloff-Coste and J. Zúñiga, *Refined estimates for some basic random walks on the symmetric and alternating groups*, ALEA Lat. Am. J. Probab. Math. Stat. **4** (2008), 359–392. MR MR2461789
36. Clyde H. Schoolfield, Jr., *Generating a random signed permutation with random reversals*, J. Theoret. Probab. **18** (2005), no. 4, 911–931. MR MR2289938 (2007m:60026)
37. Richard Stong, *Random walks on the two extra-special groups*, Tech. report, Dept. of Mathematics, Rice University, 1994.
38. ———, *Random walks on the groups of upper triangular matrices*, Ann. Probab. **23** (1995), no. 4, 1939–1949. MR MR1379174 (97c:60172)
39. Michio Suzuki, *Group theory. II*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 248, Springer-Verlag, New York, 1986, Translated from the Japanese. MR MR815926 (87e:20001)
40. Nathaniel Thiem and Vidya Venkateswaran, *Restricting supercharacters of the finite group of unipotent uppertriangular matrices*, Electron. J. Combin. **16** (2009), no. 1, Research Paper 23, 32. MR MR2482091
41. David Bruce Wilson, *Mixing times of Lozenge tiling and card shuffling Markov chains*, Ann. Appl. Probab. **14** (2004), no. 1, 274–325. MR MR2023023 (2004m:60155)
42. N. Yan, *Representation theory of the finite unipotent linear group*, Ph.D. thesis, Dept. of Mathematics, University of Pennsylvania, 2001.
43. Maria Zack, *Measuring randomness and evaluating random number generators using the finite Heisenberg group*, Limit theorems in probability and statistics (Pécs, 1989), Colloq. Math. Soc. János Bolyai, vol. 57, North-Holland, Amsterdam, 1990, pp. 537–544. MR MR1116809 (92m:65013)

Departments of Mathematics and Statistics, Stanford University, Stanford, California 94305